Communications of the Association for Information Systems

Volume 45 Article 7

8-2019

Information Systems Research: Thinking Outside the Basket and Beyond the Journal

Brian Fitzgerald
Lero, brian.fitzgerald@lero.ie

Alan R. Dennis
Indiana University

Juyoung An
Yonsei University

Satoshi Tsutsui
Indiana University

Rishikesh C. Muchhala
PricewaterhouseCoopers LLP

Recommended Citation

Fitzgerald, B., Dennis, A. R., An, J., Tsutsui, S., & Muchhala, R. C. (2019). Information Systems Research: Thinking Outside the Basket and Beyond the Journal. Communications of the Association for Information Systems, 45, pp-pp. https://doi.org/10.17705/1CAIS.04507

Communications of the Association for Information Systems

Debate DOI: 10.17705/1CAIS.04507 ISSN: 1529-3181

Information Systems Research: Thinking Outside the Basket and Beyond the Journal

Brian Fitzgerald
Lero
University of Limerick
brian.fitzgerald@lero.ie
Ireland

Alan R. Dennis
Kelley School of Business
Indiana University
USA

Juyoung An
Department of Library and Information Science
Yonsei University
Korea

Satoshi Tsutsui
School of Informatics, Computing and Engineering
Indiana University
USA

Rishikesh C. Muchhala
PricewaterhouseCoopers LLP
USA

Abstract:

Information systems (IS) researchers have long discussed research impact and journal rankings. We believe that any measure of impact should pass the same fundamental tests that we apply to our own research: validity and reliability. In this paper, we examine the impact of journals in the AIS Senior Scholars' basket of eight journals, three close contenders (i.e., journals that researchers frequently suggest for inclusion in the basket), and six randomly selected IS journals (from the Web of Science list) using a variety of traditional measures (e.g., journal impact factor) and newer measures (e.g., PageRank). Based on the results, we make three rather unpleasant and likely contentious conclusions. First, journal impact factor and other traditional mean-based measures do not represent valid measures so we conclude that one should not use them to measure journal quality. Second, the journal basket does not reliably measure quality, so we conclude that it one should not use it to measure journal quality. Third, the journal in which a paper appears does not reliably measure the paper's quality, so we conclude that one should not use the number of papers an author has published in certain journals as a criterion for promotion and tenure assessments. We believe that the best way forward involves focusing on paper-level and not journal-level measures. We offer some suggestions, but we fundamentally conclude that we do not know enough to make good recommendations, so we need more research on paper-level measures. We believe that these issues pertain to many disciplines and not just the IS discipline and that we need to take the lead in doing research to identify valid and reliable measures for assessing research impact.

Keywords: Journal Quality, Senior Scholars, Journal Basket, Paper-level Metrics, Impact Factor.

1 Introduction

The categorization of journals and their impact is a contentious subject in many disciplines, such as management (Mingers & Wilmott 2013; Singh, Haddad, & Chow, 2007), mathematics (Rousseau, 1988), psychology (Smart & Elton 1982), and various medical disciplines (e.g., Hannson, 1995; Opthof 1997; Saha, Saint, & Christakis, 2003). The issue has also exercised IS academics (e.g., Cuellar, Truex, & Takeda, 2016b; Gillenson & Stutz, 1991; Hamilton & Ives, 1980; Katerattanakul, Razi, Han, & Kam, 2005; Lowry et al., 2013; Peffers & Ya, 2003; Stewart, & Cotton, 2018; Valacich, Fuller, Schneider, & Dennis, 2006). In 2006, the AIS College of Senior Scholars created a basket of eight journals to signal which information systems (IS) journals had the highest quality (Currie et al., 2016). The Senior Scholars reviewed the basket in 2011 and 2016 and recommended no changes (Currie et al., 2016).

This raises the obvious question: what constitutes “high-quality” research? We believe that high-quality research influences subsequent research—it influences other scholars in their work (Trieschmann, Dennis, Northcraft, & Niemi, 2000). We agree that research may also have other impacts (e.g., practice, grand challenges), but we focus only on impact on subsequent research in this paper. While many possible measures to assess research impact exist (as we discuss later in this paper), researchers have commonly measured it via citations. Indeed, the fact that someone cited a paper almost always clearly and unmistakably signifies that they found it useful in their own research. In making the case for the importance of citations, Garfield (1979) showed the correlation between citations and future Nobel Prize winners.

Historically, we have used as journal’s quality as a proxy for a paper’s quality (Swanson, 2004; Trieschmann et al., 2000). Hence, we often consider a paper in a high-quality journal to be a high-quality paper. However, this view lacks logic and exemplifies the classic error of affirming the consequent[1]. Thus, while the conclusion may still be true, one has no logical way to draw a conclusion about a paper’s quality by considering the journal that publishes it.

Likewise, we often consider papers in lower-quality journals to be lower-quality papers, which we also cannot conclude (Webster & Watson, 2002). Empirical research has shown that researchers can substantially misclassify papers in that high-quality papers can appear beyond high-quality journals. In studying management journals, Singh et al. (2007) found that journals outside the commonly accepted top-tier “basket of five” journals in the management discipline published more than 75 percent of “top” papers (i.e., those with higher citation counts than the median number of citations for the dataset). Does this same pattern hold for the AIS Senior Scholars’ basket of eight journals?

In this paper, we examine the impact of papers in 17 IS journals: the basket of eight, three other journals that researchers have proposed the basket should include (*DSS*, *I&M*, *I&O*) (Currie et al., 2016), and six randomly selected journals that we drew from the Web of Science (*ECRA*, *EIS*, *IJEC*, *IJHCS*, *IJIM*, and *KBS*) (see Table 1). We use a variety of empirical impact measures to evaluate and rank this overall set of journals. Our study involves several aspects.

- First, while the basket, based on the expert opinion of a globally distributed set of AIS Senior Scholars, would seem to robustly categorize quality, our data do not validate this categorization. We used various journal-level measures based on citations, but none reproduced the basket in its entirety as the top set of eight IS journals.
- Second, given the widely reported flaws in journal impact factor (JIF), we examined a newer, iterative journal-ranking method, PageRank, to rank IS journals. Again, we did not reproduce the basket of eight journals.
- Third, given the difficulties that arise when using traditional journal citation measures (which depend on the arithmetic mean) when dealing with highly skewed data, we analyzed the journal set using median citation values. This analysis emphasizes the problems with mean analysis to the extent that a journal impact factor of 7.000 for an IS journal does not really indicate that the typical paper in the journal has attracted seven citations; a better estimate is about four citations.
- Fourth, researchers have suggested that high-quality papers often appear outside the “best” journals (Singh et al., 2007). We examined the 50 most cited *JMIS* papers in our data set (excluding special issues) and found that *MISQ* (or another journal; details below) had rejected 47 percent, yet these papers had citation rates similar to a typical *MISQ* paper (both those that

[1] Formal logic uses the form of “if p then q”. For example, if p (high-quality paper), then q (published in high-quality journal). Affirming the consequent refers to the error in which one sees q and concludes p (i.e., using the journal as an indicator of paper quality).

MISQ rejected and that it did not). This finding suggests that a paper itself represents the key determinant for high citations and that papers in journals that researchers perceive as having the highest rankings do not necessarily receive high citations.

- Finally, recognizing the fact that citations represent a "lagging indicator" of research impact (i.e., they arise after the activity of interest has occurred (O'Sullivan & Sheffrin, 2003)), we discuss the range of "leading indicators" (views, downloads mentions, recommends—activities that occur before the phenomenon of interest (i.e., impact)) that arise in paper-level metrics and altmetrics and propose an agenda for IS research to lead in this regard.

Table 1. Journals Selected for Analysis

Source	Journal
AIS Senior Scholars' basket of eight journals (in alphabetic order)	*European Journal of Information Systems* (*EJIS*)
	Information Systems Journal (*ISJ*)
	Information Systems Research (*ISR*)
	Journal of the Association for Information Systems (*JAIS*)
	Journal of Information Technology (*JIT*)
	Journal of Management Information Systems (*JMIS*)
	Journal of Strategic Information Systems (*JSIS*)
	MIS Quarterly (*MISQ*)
Proposed for inclusion in the basket	*Decision Support Systems* (*DSS*)
	Information and Management (I&M)
	Information & Organization (I&O)
Randomly selected from the Web of Science set of 51 IS-relevant Journals	*Electronic Commerce Research and Applications* (*ECRA*)
	Enterprise Information Systems (EIS)
	International Journal of Electronic Communication (*IJEC*)
	International Journal of Human Computer Studies (*IJHCS*)
	International Journal of Information Management (*IJIM*)
	Knowledge Based Systems (*KBS*)

The paper proceeds as follows. In Section 2, we examine previous research that has considered research impact and journal rankings in the IS discipline. In Section 3, we analyze our set of 17 IS journals using various traditional journal-level measures that rely on the arithmetic mean of paper citations. In Section 4, we examine two other approaches to assess impact (median and PageRank). In Section 5, we switch our focus to paper-level metrics and altmetrics. Finally, in Section 6, we conclude the paper.

2 Background

2.1 Assessing Research Quality

Evaluating the quality of research published in journals has long been an important issue for IS researchers (Gillenson & Stutz, 1991; Hamilton & Ives, 1980), and the topic continues to attract great interest today (Lowry et al., 2013; Cuellar, Takeda, Vidgen, & Truex, 2016a; Cuellar et al., 2016b; Stewart, & Cotton, 2018).

One can evaluate journal quality in many ways, though two common approaches include opinion-based journal lists (whether based on top scholars' suggestions or open surveys of many researchers) and citation-based journal lists (Ferrat, Gorman, Kanet, & Salisbury, 2007; Fisher, Shanks, & Lamp, 2007; Katerattanaku & Han, 2003; Peffers & Ya 2003). Some journal rankings aggregate prior rankings (e.g., Lewis, Templeton, & Luo, 2007; Rainer & Miller, 2005). No one way to measure quality is better or worse than another; each has its strengths and limitations. Likewise, not all faculty believe that either opinion-based or citation-based lists are appropriate (Willcocks, Whitely, & Avgerou, 2008) since such lists overrepresent North American journals and often overlook new journals (Fisher et al. 2007; Katerattanaku & Han, 2003; Willcocks et al., 2008).

Many opinion-based journal lists exist, such as the AIS Senior Scholars' basket of eight journals (Currie et al., 2016). Several countries or groups of countries have developed their own ranking lists (Fisher et al., 2007). Some pertain to specific regions, such as the Nordic List (a cooperation between Denmark, Finland, and Norway[2]) or the Australian Business Deans Council (ABDC) Journal Quality List[3]. Others pertain to countries around the world, such as the *Financial Times* top 50 journals[4] or the Academic Journal Guide (AJG)[5].

Many researchers believe citation-based journal lists to be more "objective" than opinion-based lists (Gallivan, 2011; Mingers & Willmott, 2013). Often, they view journal rankings that rely on opinions as "subjective" and argue that familiarity, anchoring, and selection biases distort them (Polites, et al., 2009), whereas they see journal rankings that rely on citations as more "objective" because citations reflect actual use and not espoused value. However, one can also manipulate citations (Polites & Watson, 2009).

Researchers have long used citations as a measure of quality. Citations almost always indicate research's positive utility: a citation means that a researcher found something useful in a paper for their own research. Cuellar et al. (2016b) suggest that citations reflect a "fitness for use" and argue that a study is neither true nor false and that a paper's fate depends on how subsequent research adopts its arguments. Researchers adopt and reuse research that they see as true in the form of citations. Thus, citations represent a good way to measure research that researchers widely see as true and, thus, indicate quality (Cuellar et al., 2016b; Garfield, 2006; Lowry et al., 2013).

Citation-based lists are not perfect because disciplines can influence citations: some disciplines receive more citations than others (Gallivan, 2011). Citations can also reinforce themselves (Mingers & Willmott, 2013) for two reasons. First, researchers are more likely to cite papers in prestigious journals (according to citation-based measures) simply because they appear in more prestigious journals (Hamilton & Ives, 1980). Second, citation measures influence faculty opinions, so citation data often shape opinion-based lists. For these reasons, citation-based measures are arguably the most important influencer of journal rankings (Mingers & Willmott, 2013).

Regardless of whether they adopt an opinion- or citation-based approach, journal lists typically come in two forms: 1) ordinal lists, which give each journal an assigned position from 1 to n, and 2) strata lists, which group journals into different strata that indicate different levels of quality (Cuellar et al., 2016b; Gillenson & Stutz, 1991; Hamilton & Ives, 1980). The Australian Business Deans Council (ABDC) list and Senior Scholar's basket exemplify a strata list, while Lowry et al. (2013) list journals based on both types: they identify two journal strata and then use an ordinal ranking for the top stratum (see also Stewart & Cotton, 2018).

Several studies have analyzed journal lists (both opinion-based and citation-based) over time, and the results have not always been consistent. Katerattanaku and Han (2003) found few significant differences in citations among the eight journals in the Senior Scholar's basket, although *MISQ* had significantly more citations than other journals depending on the measure. Katerattanakul et al. (2005) found significant differences in journal rankings that opinion-based methods created versus rankings that citation-based methods created. Lewis et al. (2007) examined papers that ranked IS journals and found significant correlation and consistency among them; they concluded this body of research exhibited good content, convergent and discriminant validity, and reliability, so we have reason to believe that past journal-ranking papers can help one determine high-quality journals. Xiao, Cheung, and Thadani (2011) found that *MISQ* had significantly more citations than *ISR* and *JMIS*. Lowry et al. (2013) examined the basket and13 other journals from the Web of Science and concluded that *MISQ*, *ISR,* and *JMIS* comprise the top tier and that *DSS*, *EJIS*, *I&M*, *IJEC*, *ISJ*, *JAIS*, *JIT,* and *JSIS* form a second tier of top journals.

2.2 The Journal as a Measure of Research Quality

The predominant method to evaluate research quality has been to focus on the journal (Cuellar et al., 2016b; Swanson, 2004; Trieschmann et al., 2000). All of the studies above used the journal as the central construct of interest. Opinion-based quality approaches solicit faculty opinions on journal quality, not paper quality, but faculty typically base their opinions on the papers they have read and, thus, generalize their opinions

[2] https://dbh.nsd.uib.no/publiseringskanaler/Forside.action?request_locale=en
[3] http://www.abdc.edu.au/master-journal-list.php
[4] https://www.ft.com/content/3405a512-5cbb-11e1-8f1f-00144feabdc0
[5] https://charteredabs.org/academic-journal-guide-2015/

about a set of papers to journals as a whole. Also, not all faculty expend equal effort in reading papers in all journals, so faculty opinions often do not represent a complete and reliable measure of journal quality. Likewise, citation-based quality approaches measure citations for individual papers but aggregate them into one measure at the journal level.

In other words, we generalize from measures of papers quality to produce a measure of journal quality. We also do the same process in reverse: we generalize from journal quality to paper quality: we often perceive papers in high-quality journals to be high-quality papers and papers in low-quality journals to be low-quality papers (Fisher et al. 2007; Cuellar et al., 2016b).

Using the journal as the measure of research quality presents two fundamental problems: 1) a theoretical problem and 2) an empirical problem. First, using the journal as a measure of research quality shifts the focus (and resources) to the journal as what creates value and away from its actual creators: the paper and its authors (Cuellar et al., 2016b). The journal primarily has a selection or curation function; in some cases, the journal adds value to the paper during the review process. By focusing on the journal and not the paper, we measure quality in the wrong place.

Second, one can use the journal as the measure of research quality validly only when the papers in a given journal are of an equivalent quality. From analyzing papers in the management discipline, Singh et al. (2007) found that only about one quarter of the most highly cited papers appeared in the top five journals. In a similar vein, researchers have suggested that papers that top journals reject by definition cannot achieve citations. To elucidate this issue, we investigated the top 50 highly cited papers in *JMIS* and surveyed the authors as to whether these *MISQ* or *ISR* had previously rejected them. We report our findings at the end of Section 4.1 below.

In the IS discipline, Cuellar et al. (2016b) examined the uniformity of journals in different strata (rather than papers in journals). They examined the strata in the United Kingdom's AJG list, the Australian ABCD list, the AIS Senior Scholar's basket, and the *Financial Times* list and found significant differences in citations between strata. Specifically, they found journals in higher strata had more citations than journals in lower strata. However, they also found that journals in the same stratum usually had significantly different citation counts, which indicates that not all journals in a given stratum were of an equivalent quality. Perhaps more importantly, they found that about 98 percent of papers in papers in the top strata of the lists did not qualify for inclusion in the top strata; they should have been in journals in lower strata. Thus, they conclude that, although they found differences in quality among strata, the strata do not reliably indicate research quality. They argue that one should not measure individual scholars' research productivity by counting papers in journals in different strata. Instead, they argue that research productivity should be a composite measure of scholarly capital that has three dimensions: impact (e.g., h-index), position in co-author networks (e.g., centrality), and publications in specific journals (Cuellar et al., 2016a).

Regardless of whether one agrees that journal lists are useful tools or impediments to good research (Willcocks et al., 2008), journal lists clearly matter. They shape faculty research behavior (Mingers & Willmott, 2013; Willcocks et al., 2008), and output as measured by the number of publications in ranked journals has real consequences for individuals, departments, and institutions (Mingers & Willmott, 2013).

These issues raise two important questions: is the journal an appropriate measure of research impact in terms of validity and reliability (Campbell & Stanley, 1963; Cronbach & Meehl, 1955), and is the journal a reasonable measure of the quality of the papers it contains? That is, can we use the aggregate journal-level measure of quality (as assessed using previously published papers) as a reasonable proxy to estimate newly published papers' quality? More specifically, can we generalize journal-level aggregate measures to the individual papers that each journal contains?

2.3 Selecting the Data Set

Lewis et al. (2007) estimated that more than 500 journals publish IS-related papers, though not all publish purely IS papers. The Web of Science journal database lists 51 IS-relevant journals. Clearly, with such a number of journals, researchers will perceive some measure of ranking journal quality as useful. In this context, using domain experts' expertise has significant advantages. The College of Senior Scholars provides one influential list that ranks IS journals' quality: the AIS Senior Scholars' basket of eight journals. Researchers achieve "Senior Scholar" status through having significant research leadership in the IS discipline. Given their longstanding background and status in the discipline, the AIS Senior Scholars have a good position to assess the quality of IS journals (Cuellar et al., 2016b).

In order to assess the quality of papers and journals, we needed to select a set to analyze. We begin with a set of journals that experts selected: the AIS Senior Scholars' basket. Note that the basket grouping has evolved over time: it began with just two journals initially (*ISR* and *MISQ*) before the Senior Scholars added four more journals (*EJIS*, *ISJ*, *JAIS*, and *JMIS*) to make the "basket of six" in 2006 and a final two (*JIT* and *JSIS*) after that. The journals in the basket represent the "high-quality" journals in the IS discipline according to the Senior Scholars and many other scholars in our discipline. Thus, we used these eight journals in our analyses.

When the Senior Scholars reassessed the basket in 2016, many surveyed IS researchers proposed including three additional journals (i.e., *DSS*, *I&M*, *I&O*) in an expanded basket (Curry et al., 2016). Thus, we included these journals in our analyses as well given that they have attained at least somewhat wide recognition as "high-quality" journals.

We also needed to select a set of journals that represented journals not in the two "high-quality" sets we already selected. Therefore, we chose six journals at random from the list of IS-relevant journals in the WOS database (*ECRA*, *EIS*, *IJEC*, *IJHCS*, *IJIM*, *KBS*).

As such, we obtained 17 journals to include in our analyses (see Table 1). Note that we excluded boundary journals that publish IS research but not exclusively from our analyses (e.g., *Communications of the ACM*, *Decision Sciences*, and *IEEE Software*). We agree that such boundary journals are good and appropriate places to publish IS research intended to reach a broader audience. However, we focus on IS research, not research in other disciplines, and including boundary journals would conflate IS research with research in other disciplines.

3 Traditional Journal-level Measures

Many journal citation measures to rank journals exist, such as journal impact factor, journal impact factor excluding self-citations, five-year journal impact factor, immediacy, and cited half-life. In this section, we explain each measure and present the journal rankings if one applied that particular measure. Since these measures tend to vary considerably from year to year, we averaged all measures over a five-year period (from 2011 to 2015). We also used a five-year window to ensure that papers had a sufficient "window of opportunity" in which to accrue citations (Campanario, 2011; Singh et al., 2007). We also needed to strike a balance in relation to the discipline's "currency". The IS discipline is quite dynamic as its technology focus ensures that new concepts frequently emerge. If we chose too long a time window, we ran the danger that it would not reflect the discipline's current state. A five-year window also serves as the basis for the Eigenfactor Project's measures whose creators designed to be more reliable than traditional measures[6].

3.1 Journal Impact Factor (JIF) 2011-2015

The journal impact factor (JIF) is the number of citations in the current year to papers published in the previous two years divided by the total number of papers. For example, *MISQ*'s JIF in 2015—5.384—comes from an overall total of 603 citations in 2015 and the 112 citable papers (i.e., research papers excluding editorials) that the journal published in 2014 and 2013 (603 divided by 112 = 5.384). Table 2 presents the JIF values for the 17 journals in our dataset.

JIF values can significantly fluctuate from year to year, so we calculated the average JIF in the 2011-2015 period. Using this ranking measure, we found that *EIS* occupied the number one spot. However, note that we did not have complete data for *EIS* for all five years since the Web of Science suppressed the journal in 2013 and 2014. Thomson Reuters states that "suppressed journals represent extreme outliers in citation behavior"[7]. The JIF for *EIS* increased almost threefold to 9.256 between 2011 and 2012; thus, the Web of Science suppressed the journal in 2013 and 2014 and did not record a JIF for it in these years.

Table 2 shows that *MISQ* occupied the number two spot after a journal we randomly selected (*EIS*). *JIT* occupied the third spot and *ISR* the fifth. Also, the basket of eight would now include *EIS*, *DSS*, *KBS*, and *IJIM*. Of the original basket of eight, *EJIS*, *JMIS*, *ISJ,* and *JAIS* made way for four newcomers, two of which we randomly selected. In a general linear model (GLM) analysis, we found significant differences in JIF among the 17 journals ($F(16,83) = 1068.67$, $p = .000$). In a post hoc Tukey analysis, we found four groupings, which Table 2 shows. The Tukey Group shows journals that resemble or differ from each other depending

[6] http://www.eigenfactor.org/

[7] One can access an overview of the suppression method at http://wokinfo.com/media/pdf/jcr-suppression.pdf

on their JIF. For example, *EIS* and *MISQ* were in the same group (group one), which means their JIF did not significantly differ from each other but did significantly differ from the journals in all other groups. *JIT* was in groups one and two, which means it did not significantly differ from the journals in those groups but did differ from the journals in groups three and four. We found no significant difference in JIF between the traditional basket journals and non-basket journals ($F(1,98) = 2.52$, $p = .115$).

Table 2. Journal Impact Factor Rankings 2011-2015[8]

Journal	Avg. JIF 2011-2015	Tukey group	Avg. items published per year
EIS	5.070	1	28
MISQ	5.041	1	53
JIT	3.788	1, 2	20
KBS	3.171	2, 3	260
ISR	2.392	2, 3, 4	54
DSS	2.168	3, 4	172
JSIS	2.163	3, 4	21
IJIM	1.986	3, 4	72
EJIS	1.963	3, 4	38
I&M	1.939	3, 4	75
JMIS	1.939	3, 4	36
ISJ	1.814	3, 4	22
IJEC	1.770	3, 4	18
ECRA	1.575	4	45
I&O	1.546	4	14
JAIS	1.506	4	29
IJHCS	1.304	4	68

As in many disciplines, IS researchers and journal editors seem to have developed an unhealthy preoccupation with the journal impact factor (JIF). Researchers have well documented JIF's limitations (e.g., Perneger, 2010; Vanday, 2012). The JIF is a simple mean that purportedly represents the average citations for the typical papers. In reality, the citations almost always have a highly skewed spread in that a small number of papers receive a high number of citations while the remaining papers receive few or no citations at all. Seglen (1992) found that typically 15 percent of the papers in a journal account for more than 50 percent of its citations. The two-year window that the JIF considers is also problematic as citations can accumulate naturally over a much longer period. Indeed, we discuss the merit of a five-year window in the introduction to Section 3 above. The JIF also uses three decimal places. While these decimal places help avoid tied scores among journals from occurring, they lack true accuracy—what does one thousandth of a citation actually represent? Also, citation patterns vary significantly across disciplines as certain disciplines can have more or fewer citing researchers.

Despite these shortcomings, the academic community frequently uses the JIF as a proxy for quality. Researchers include it in their résumés. Recruitment and promotion panels and funding agencies consider it when making decisions. The JIF is also somewhat circular and self-fulfilling since researchers will view journals with high impact factors as more prestigious and, thus, be more likely to cite such journals—something that Perneger (2010) has confirmed. He investigated various situations where several journals published identical papers (white papers or consensus reports) and found an almost perfect correlation between citations received and the impact factor of the journal in which the papers appeared; that is, a

[8] Note that, in Table 2 and subsequent tables, we identify the original basket of eight journals with light shading to help readers easily identify them. The empty row indicates the top-eight journal cut-off point. We had incomplete data for *EIS* since the Web of Science "suppressed" it in 2013 and 2014 as it had a substantially higher JIF in these years than in other years. Journals in different Tukey groups had significantly different JIF.

paper in a journal with a higher JIF received correspondingly more citations than the same paper in a journal with a lower JIF.

3.2 Journal Impact Factor Excluding Self-citations (JIF-SC) 2011-2015

Table 3 presents the ranking using the journal impact factor excluding self-citations (JIF-SC) measure, which we averaged over the 2011-2015 period. Self-citation in a journal context refers to journal papers' citing other papers in the same journal. *MISQ* occupied the top position when we removed self-citations, and *JIT*, *ISR,* and *ISJ* also remained in the top eight along with *EIS* (with the usual caveat), *KBS*, *DSS,* and *I&M*. However, this ranking would exclude four original basket journals (*EJIS*, *JSIS*, *JMIS*, and *JAIS*) from the basket. In a GLM analysis, we found significant differences in JIF-SC among the journals ($F(16,83) = 1059.16$, $p = .000$). In a post hoc Tukey analysis, we found two groupings, which Table 3 shows. The traditional basket journals had significantly higher JIF-SC than non-basket journals ($F(1,98) = 7.50$, $p = .007$).

Self-citations (which arise when a journal's papers cite other papers in the journal to inflate the journal's citations) involve some controversy since one can see them as a mechanism to artificially increase a journal's JIF. However, good journals publish good research, which other papers in that journal will inevitably cite. In studying over 1.5 million papers, King, Correll, Jacquet, Bergstrom, and West (2016) found that 9.5 percent of citations were self-citations (i.e., to other papers in the same journal). As such, that figure seems to represent a reasonable benchmark[9]. For example, when we exclude self-citations in the prestigious journals *Nature* and *Science*, their JIF reduces by two percent and one percent, respectively. As for IS journals, their JIF reduced much more severely (see the final column in Table 3). The reduction ranged from under 10 percent for *JIT*, *ISJ*, *IJHC*, *JAIS*, and *MISQ* to over 20 percent for *JMIS*, *JSIS*, *IJIM*, *DSS*, *EIS*, and *I&O* and over 30 percent for *IJEC* and *KBS.*

Table 3. Journal Impact Factor (Excluding Self-citations) 2011-2015

Journal	Avg JIF-SC 2011-15	Tukey Group	% Reduction from JIF
MISQ	4.568	1	9%
EIS	3.939	1	22%
JIT	3.550	1	6%
ISR	2.137	2	11%
KBS	1.962	2	38%
DSS	1.690	2	22%
I&M	1.672	2	14%
ISJ	1.672	2	8%
EJIS	1.668	2	15%
JSIS	1.616	2	25%
IJIM	1.531	2	23%
JMIS	1.400	2	28%
JAIS	1.390	2	8%
ECRA	1.312	2	17%
I&O	1.215	2	21%
IJHCS	1.203	2	8%
IJEC	1.160	2	34%

[9] The more newsworthy aspect in King et al.'s (2016) study involved gender. Over the entire period the authors studied, men self-cited their own work 56 percent more often than women, and the trend appears to have worsened over time: for just the past 20 years, men self-cited their own work 70 percent more often than women. The authors found this result across all disciplines, which presumably included IS.

3.3 Five-year Journal Impact Factor (5Y-JIF)

Some researchers have criticized the regular JIF for only including citations in the previous two years since, given long review cycles, papers simply do not have enough time to accrue an indicative number of citations. Thus, the five-year JIF measure captures a longer period to allow such citations to accrue. Table 4 presents the journal ranking using the five-year JIF. *MISQ*, *JIT,* and *ISR* occupied the top three places with *JMIS* and *JSIS* also in the top eight. *ISJ*, *JAIS*, and *EJIS* dropped out of the top eight.

3.4 Immediacy

In contrast to expanding the JIF timeframe to five years, one can use the immediacy measure: the number of citations a paper attracts in its year of publication. This measure rests on the rationale that, in a technology-oriented discipline such as IS, research can become obsolete quite quickly and researchers will find the latest research more noteworthy. Table 4 also presents the journal ranking using immediacy. Here, *ISJ* topped the list with *MISQ* in third place. Six of the basket journals retained their place (*ISJ*, *MISQ*, *JIT*, *JSIS*, *ISR,* and *EJIS*), while *JAIS* and *JMIS* made way for *EIS* and *KBS*.

3.5 Cited Half-life

Cited half-life (CHL) falls at the other end of the spectrum from the immediacy index (see Table 4). CHL represents the median paper publication date—half of the cited papers were published before this time, half were published afterwards (i.e., a cited half-life of 5 in 2015 indicates that half the citations in 2015 cite papers published more than five years earlier). The CHL measure reflects a journal's longevity in some respects and has "> 10 years" as its upper-level cut-off (e.g., *Nature* and *Science* have a > 10 cited half-life). In the basket of eight, *MISQ* and *ISR* fell into this category. The next six journals were *JMIS*, *JSIS*, *IJHCS*, *I&M*, *IJEC,* and *ISJ*. Of the original basket of eight, this measure retained five (*ISR*, *MISQ*, *JMIS*, *JSIS,* and *ISJ*). *EJIS*, *JIT,* and *JAIS* made way for *IJHCS*, I&M, and *IJEC*.

Table 4. Five-year JIF, Immediacy, and Cited Half-life

Journal	Avg. five-year JIF 2011-2015	Journal	Avg. immediacy 2011-2015	Journal	Avg. CHL 2011-15
MISQ	8.226	*ISJ*	1.213	*MISQ*	>10
JIT	4.651	*EIS*	1.160	*ISR*	>10
ISR	3.963	*MISQ*	0.856	*JMIS*	9.6
EIS	3.887	*JIT*	0.802	*JSIS*	8.8
I&M	3.329	*JSIS*	0.590	*IJHCS*	8.7
JMIS	3.175	*KBS*	0.589	*I&M*	8.7
KBS	2.940	*ISR*	0.373	*IJEC*	8.6
JSIS	2.887	*EJIS*	0.355	*ISJ*	7.3
DSS	2.845	*I&O*	0.305	*EJIS*	6.9
IJEC	2.778	*IJHCS*	0.301	*JIT*	6.7
ISJ	2.735	*IJIM*	0.279	*I&O*	6.6
JAIS	2.727	*DSS*	0.232	*JAIS*	6.2
EJIS	2.543	*JAIS*	0.227	*DSS*	6.1
IJIM	2.316	*JMIS*	0.215	*IJIM*	5.6
I&O	2.184[10]	*I&M*	0.175	*ECRA*	4.9
ECRA	2.112	*ECRA*	0.162	*KBS*	3.2
IJHCS	1.968	*IJEC*	0.158	*EIS*	3.0

[10] Note: five-year JIF was not available in Web of Science for *I&O* in 2011 and 2012.

3.6 Summary of Traditional Journal-level Measures

Table 5 summarizes the rankings of individual journals for each of the measures above. We draw two conclusions from this table.

Table 5. Summary of Traditional Citation-based Rankings

Journal	JIF	JOF-SC	5Y-JIF	Immediacy	CHL	Rank Sum
MISQ	2	1	1	3	1	8
ISR	5	4	3	7	1	20
JIT	3	3	2	4	10	22
EIS	1	2	4	2	17	26
JSIS	7	10	8	5	4	34
KBS	4	5	7	6	16	38
ISJ	12	8	11	1	8	40
I&M	10	7	5	15	6	43
DSS	6	6	9	12	13	46
JMIS	11	12	6	14	3	46
EJIS	9	9	13	8	9	48
IJIM	8	11	14	11	14	58
IJHCS	17	16	16	10	5	64
IJEC	13	17	10	17	7	64
I&O	15	15	15	9	11	65
JAIS	16	13	12	13	12	66
ECRA	14	14	15	16	15	74

First, none of the empirical citation measures reproduced the AIS Senior Scholars' basket of eight ranking. *MISQ* and ISR consistently placed among the top eight across all traditional measures, and *JIT* and *JSIS* performed well on four out of the five measures, but the consistency ends there; no other basket journal consistently placed in the top eight. In fact, two randomly selected journals (*EIS* and *KBS*) placed more consistently in the top eight than four of the basket journals.

Second, we found considerable variation in rankings depending on the measure that we used. With the possible exception of *MISQ*, we found little stability. The different measures produced different conclusions about the relative quality of the basket journals, the three journals on the cusp, and six randomly selected journals. In fact, at least one (sometimes two) randomly selected journals made it into the top eight depending on the measure. Two randomly selected journals commonly appeared (*EIS* and *KBS*), but three other random journals (*IJEC*, *IJHCS*, *IJIM*) popped in and out depending on the measure.

We find this lack of stability troubling—even with measures that share a common base but lack stability (i.e., the three forms of the JIF that share a foundation on mean citations), which we believe raises serious reliability and validity concerns. In Section 4, we investigate alternative citation metrics based on median values and also an iterative mechanism based on the PageRank algorithm.

4 Non-mean-based Journal-ranking Measures

Researchers developed most journal-quality measures many decades ago (Garfield, 2006), and they all suffer from one important limitation: they use some form of arithmetic mean. The arithmetic mean measures central tendency well with data that have a normal underlying distribution, but it leads to bias with non-normally distributed data (Pagano & Gauvreau, 2000). Unfortunately, all these measures rely on citations, which lack normally distribution—they follow a power law distribution, which differs from normal distribution (Brzezinski, 2015). Citations have a long tail in that most papers have few citations, while few have many. However, the traditional measures above do not capture this distribution well. Over the intervening years, researchers have suggested several newer approaches that do not use the arithmetic mean. One such approach, the median citation measure, uses median rather than mean. Another approach, PageRank, uses incoming and outgoing links to and from webpages.

4.1 Median Citation Measure

To derive our dataset for this analysis, we examined the citations for the 6,311 papers in our set of 17 journals for a five-year period (omitting the year of publication) from 2001-2010. For example, for papers published in 2010, we examined citations during 2011-2015; for papers published in 2009, we examined citations during 2010-2014. Table 6 shows the descriptive statistics for the citation data by journal.

Researchers have established that journal citations have a somewhat skewed distribution; however, we did not know to what extent. Researchers usually consider values for skewness and kurtosis between -2 and +2 as acceptable in order to use statistics that assume a normal distribution (Field, 2009; George & Mallery, 2010; Gravetter & Wallnau, 2014). Only one journal met these bounds (*EIS*); all the others had a non-normal citation distribution (some by one or two orders of magnitude), which means we cannot use statistics that assume a normal distribution.

Perhaps more importantly, the mean noticeably differed from the median number of citations. In most cases, the 95 percent confidence interval for the mean did not contain the median and vice versa. In other words, the mean and median pointed to two very different places as the central tendency. For this set of journals, the mean was 162 percent of the median on average and at about the 69th percentile in the distribution; stated differently, the median was about 62 percent of the mean, so that a five year JIF of 7.00 would suggest a typical paper receives only four citations (not seven as the JIF implies). With highly skewed distributions, the mean does not appropriately measure central tendency; instead, the median measures it better (Pagano & Gauvreau, 2000). Thus, measures that build on the mean (see the measures we discuss in Section 3) will contain bias and give invalid measures.

Each journal had a large amount of variation. Many papers received zero, one, or two citations, and a small few (i.e., top five percent) receive many. Most journals had a coefficient of variation (COV) above 1.00, which indicates highly variable data (in other words, the standard deviation exceeded the mean). The COV indicates even greater variance when one considers that, because these are count data, no value can be below zero.

Table 6. Descriptive Statistics for Number of Citations in Five Years

Journal	Median	95% CI for median	μ	95% CI for μ	%ile of μ	Std. dev.	COV	Min	Max	Skew	Kurtosis
MISQ	30	24-36	47.0	41.0-53.0	56	13.47	0.29	0	548	4.22	27.12
ISR	17	14-20	24.5	21.2-27.8	65	28.15	1.15	0	259	3.24	18.43
JMIS	10	5-15	17.8	13.2-22.3	72	47.07	2.65	0	900	16.14	299.79
I&M	9	7-11	14.5	13.0-16.0	70	18.80	1.30	0	149	3.21	13.46
JAIS	9	6-12	13.4	10.7-16.1	74	17.42	1.30	0	126	3.94	19.80
JSIS	9	7-11	12.4	10.5-14.3	63	13.12	1.06	0	112	3.19	18.52
EIS	8.5	6-11	13.1	10.8-15.4	64	10.89	0.83	1	42	1.06	0.14
ISJ	8	6-10	11.8	10.1-13.4	64	12.15	1.03	0	115	3.82	25.72
EJIS	8	7-9	11.2	10.1-12.3	65	11.09	0.99	0	92	2.82	12.00
IJHCS	7	6-8	11.1	10.1-12.1	71	14.30	1.29	0	123	3.51	16.90
DSS	7	6-8	10.9	9.78-12.0	67	18.54	1.70	0	436	12.43	256.36
ECRA	6	4-8	11.0	9.0 -13.0	73	14.26	1.30	0	82	2.88	9.61
IJIM	6	5-7	9.4	8.2-10.5	65	12.15	1.30	0	132	4.61	34.42
IJEC	5	3-7	10.5	8.5-12.5	73	15.77	1.50	0	113	3.41	15.35
KBS	5	4-6	9.3	8.3-10.3	70	13.47	1.45	0	149	4.56	33.12
JIT	5	4-6	8.6	7.1-10.0	69	12.01	1.40	0	78	3.40	13.96
I&O	5	0-10	8.7	4.0-13.5	84	14.47	1.66	0	82	4.18	18.93

Note: CI = confidence interval, μ = mean; %ile = percentile, COV = coefficient of variation.

We can consider how each journal compares to the entire sample as a whole. The first part of Table 7 compares to overall sample percentiles: the 25th percentile of the overall sample was three citations, the 50th was eight citations, the 75th was 16 citations, and the 90th was 31 citations. A journal whose citation pattern matches the sample as a whole would have percentiles close to these marks. *I&M* exemplifies a journal that closely matches the overall sample: three cites was the 24th percentile of the journal's papers—close to the 25th percentile of the overall sample; the same was true for the other benchmarks (eight citations (overall 50th percentile) was 47th percentile, 16 citations (75th overall) was 75th percentile, and 31 citations (90th overall) was 89th percentile).

By comparison, we found that *MISQ* had a high impact level: the three citation mark (25th percentile overall) was the 2nd percentile of *MISQ* papers, and 31 citations (overall 90th percentile) was only the 51st percentile for *MISQ* papers. Conversely, those journals lower in the table had a lower impact level.

Table 7 shows that some journals have published many papers (e.g., *DSS*) while others have published relatively few (e.g., *I&O*). For example, *MISQ* accounted for about 5.1 percent of all IS papers in our sample, while *DSS* accounted for about 17.5 percent. Interestingly, about 4.7 percent of all papers with citations at or above the median also appeared in *MISQ* as did about 2.5 percent of high-impact papers (i.e., at or above the 90th percentile in impact). By comparison, about 8.1 percent of all papers cited at or above the median appeared in *DSS* as did about 1.0 percent of high-impact papers.

Table 7. Relative Impact by Journal

Journal	Papers with this impact or less as a percent of the journal				Total number of papers in the journal	Total number of papers as percent of entire sample	Papers with this impact as a percent of the entire sample			
	3 cites 25th	8 cites 50th	16 cites 75th	31 cites 90th			Below 50th %	At or above 50th %	At or above 75th %	At or above 90th %
MISQ	2%	8%	20%	51%	319	5.1%	0.4%	4.7%	4.0%	2.5%
ISR	16%	33%	48%	73%	285	4.5%	1.5%	3.0%	2.3%	1.2%
JMIS	20%	43%	68%	88%	415	6.6%	2.8%	3.7%	2.1%	0.8%
I&M	24%	47%	75%	89%	591	9.4%	4.4%	5.0%	2.4%	1.0%
JSIS	24%	48%	74%	94%	180	2.9%	1.4%	1.5%	0.7%	0.2%
JAIS	21%	49%	77%	93%	160	2.5%	1.2%	1.3%	0.6%	0.2%
EIS	19%	50%	69%	92%	86	1.4%	0.7%	0.7%	0.4%	0.1%
ISJ	21%	51%	78%	95%	206	3.3%	1.6%	1.6%	0.7%	0.2%
EJIS	19%	51%	81%	95%	383	6.1%	3.1%	3.0%	1.1%	0.3%
DSS	30%	54%	82%	94%	1107	17.5%	9.4%	8.1%	3.2%	1.0%
IJHCS	28%	58%	82%	96%	753	11.9%	6.9%	5.1%	2.1%	0.5%
ECRA	30%	61%	84%	94%	199	3.2%	1.9%	1.2%	0.5%	0.2%
IJIM	31%	62%	86%	95%	416	6.6%	4.1%	2.5%	0.9%	0.3%
IJEC	40%	63%	82%	93%	234	3.7%	2.3%	1.4%	0.7%	0.3%
KBS	40%	65%	86%	94%	669	10.6%	6.9%	3.7%	1.5%	0.6%
JIT	39%	69%	90%	96%	270	4.3%	2.9%	1.3%	0.4%	0.2%
I&O	37%	79%	92%	97%	38	0.6%	0.5%	0.1%	0.0%	0.0%

Note: percentages do not always add to 100% due to rounding.

The journals with the most cited papers (90th percentile and above) were *MISQ*, *ISR*, *JMIS*, *I&M*, and *DSS*. Each of these journals contained about one percent of the most cited papers in our discipline. One can see that the basket already contains three of these journals (*MISQ*, *ISR*, *JMIS*), while researchers have suggested the other two join the basket (*DSS* and *I&M*). We randomly selected four of the next five journals that accounted for the most cited papers (*KBS*, *IJHCS*, *IJIM*, *IJEC*). As we note above, each of these journals published a different number of papers, so that about half of the papers in *MISQ* had a high number of citations, while only about six percent of papers in *DSS* had a high number of citations.

Figure 1 shows the median number of citations that a typical paper in each journal received (as a dot). The line for each journal shows the 95 percent confidence interval around the median. The figure also shows the grand median in the entire data set (eight citations in five years).

One important question involves whether papers in different journals receive a significantly different number of citations. Due to the highly skewed distributions, traditional statistical analyses are not appropriate. Instead, we did a medians test and found significant differences (χ(df = 16) = 454.62, p = .000). Cramer's V was .268, which indicates that, although the differences were significant, the effect size was small; that, is the differences may not have been material. The medians test uses the proportion of papers above the grand median versus those at or below it. Figure 2 shows the proportion for each journal along with the 95 percent confidence interval.

As we look across the citation data in these tables and figures, we draw two overall conclusions. First, a typical paper in *MISQ* clearly had more citations than a typical paper in any of the other journal in our set. Further, a typical paper in *ISR* also had more citations than a typical paper in any other journal except *MISQ,* although it had a less clear pattern. The remaining journals may or may not had significant differences between them, but the real question concerns whether these differences were material. Figure 1 shows that, over five years, papers across the journals (excluding *MISQ* and *ISR*) had a median number of citations that ranged from up to two above or three below the grand median. Does such a difference constitute a material difference in quality?

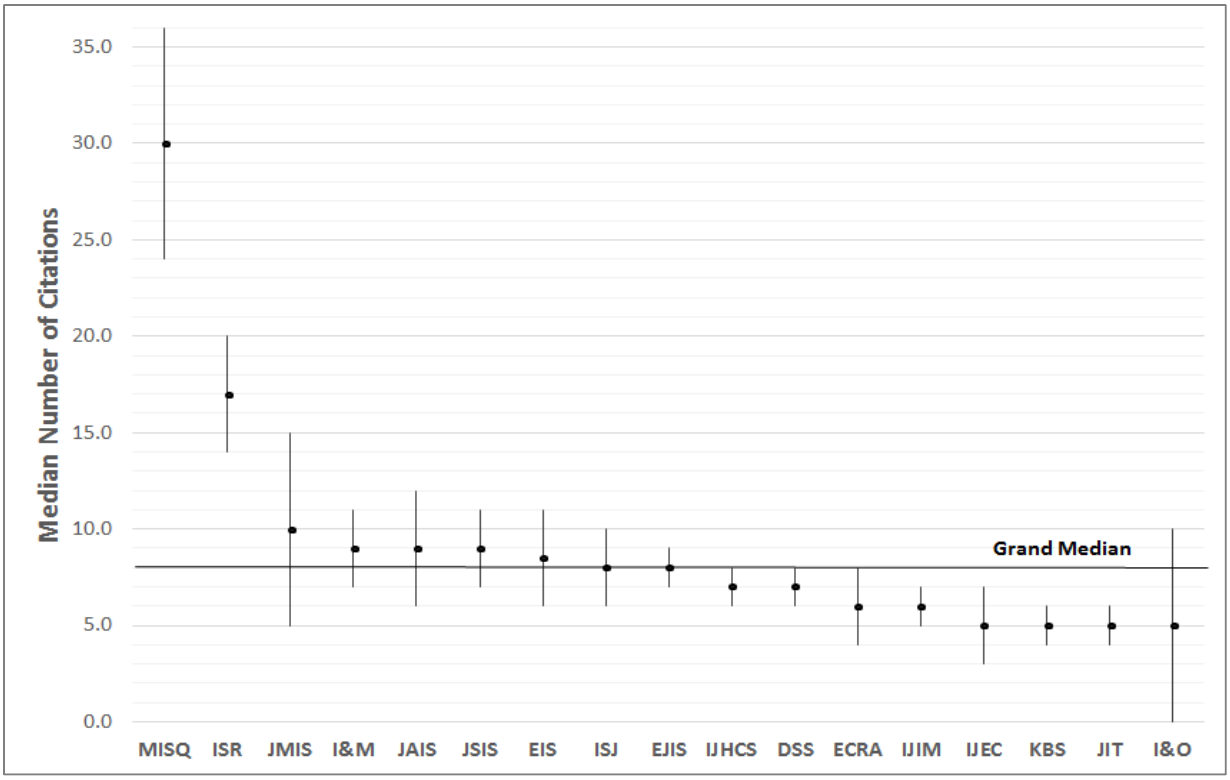

Figure 1. Median Number of Citations per Paper[11]

Second, we note the basket journals generally received more citations than non-basket journals, although we found some exceptions. We could not distinguish some non-basket journals (*I&M* and *EIS*) from the basket journals, and one basket journal (*JIT*) had similar citations to our randomly selected journals.

As we reflected on these analyses, we recalled a comment from Detmar Straub when he was editor-in-chief of *MISQ*: we can study the impact of *MISQ* papers, but we can never know the impact of a paper that *MISQ* rejected. His point involved raising concern about type II errors whereby one rejects good papers that warrant publication. We realized that, when *MISQ* rejects our papers, we always submit them to other

[11] Note: the dot shows the median and the line shows the 95 percent confidence interval around it.

journals. Could we assess the quality of papers that *MISQ* rejected by looking at papers that other journals ultimately published?

Such a study would represent a major undertaking in its own right, but we decided to examine this question in a more limited way. We selected the top 50 most cited *JMIS* papers in our data set (excluding papers that appeared in a special issue based on the logic that authors usually submit special issue papers first to the special issue). We located authors for 48 of the 50 papers and emailed them to ask if they had submitted their paper to *MISQ* (or *ISR*) prior to submitting it to *JMIS*. Authors for 44 (from the 48 possible) papers responded (92% response rate). We found that authors had submitted 20 papers (47%) to *MISQ* (or, in a very few cases, another journal) prior to submitting it to *JMIS*, authors had not submitted 23 papers to another journal first, and one author could not remember. These 20 papers had 37.9 citations on average and a median of 32.5 citations. In other words, these papers resembled a typical *MISQ* paper. The 23 papers whose authors did not submit to another journal first had 44.7 citations on average and a median of 40 citations, so they too resembled a typical *MISQ* paper.

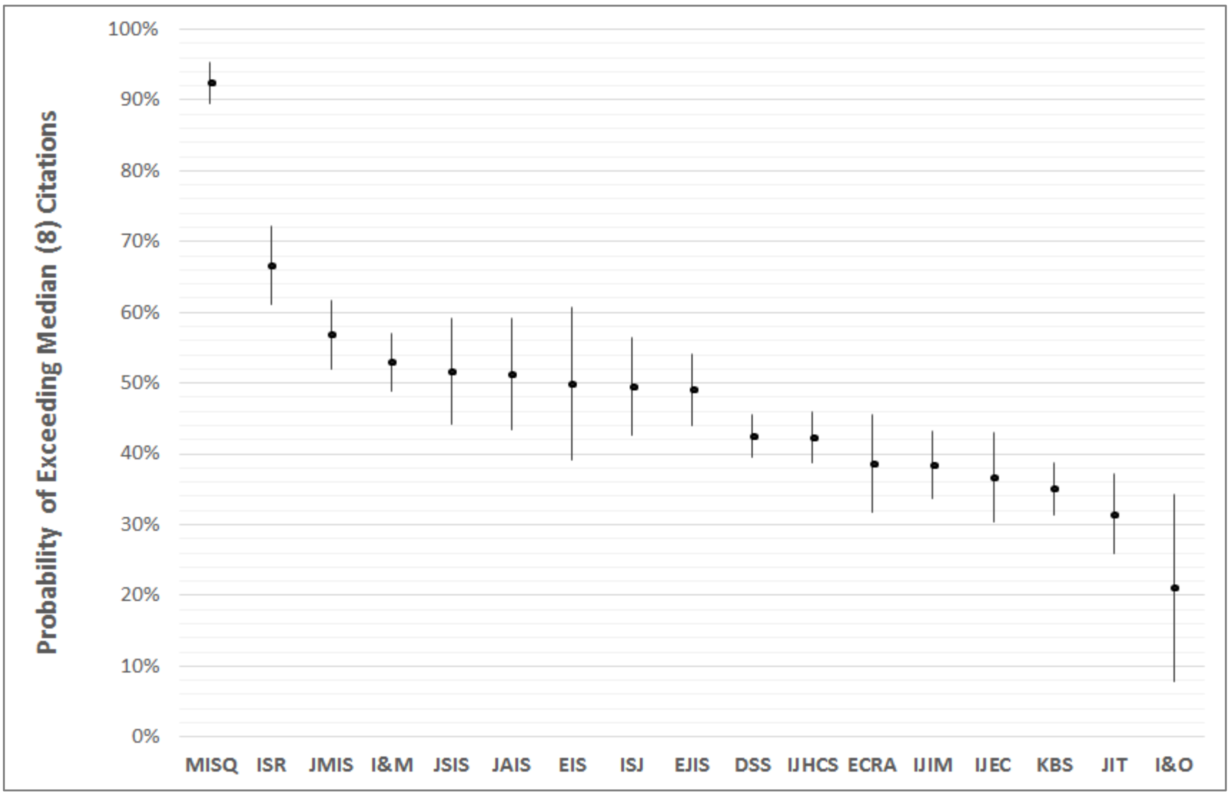

Figure 2. Probability of Exceeding the Grand Median Number of Citations per Paper[12]

4.2 PageRank Journal Comparisons

Another more recent measure, PageRank uses incoming and outgoing links to and from webpages to rank their importance (Brin & Page, 2012). Ironically, its creators found inspiration in Garfield's (1979) work on citation analysis, so it is fitting that we should apply it to citation analysis again here to rank our set of IS journals.

In our context, a webpage corresponds to a paper and the number of links corresponds to citations. In this context, PageRank relies on the key idea that a quality paper has received many citations from papers that have also received many citations while controlling for the number of citations in it and the other papers. In other words, a citation from a highly cited paper more strongly indicates quality than a citation from a paper that has never received any citations. Likewise, a citation has more importance if the citing paper cites fewer papers in total (e.g., if a paper received citations from another paper that cited 20 papers, the citation would

[12] Note: the dot shows the mean probability and the line shows the 95 percent confidence interval around it.

have more importance than if the citing paper cited 1000 papers)[13]. Citation practices differ by discipline or subject— some disciplines or topics receive significantly more citations than others (Gallivan, 2011)—so one needs to control for these practices. Therefore, PageRank controls for the number of citations that the target paper makes to other papers (and, thus, all the papers in the citation chain) as a way to control for these inherent differences in citation practices in different topic areas.

We conducted a PageRank analysis using the 51 journals that Web of Science defined as IS journals. This set includes what we consider to be IS journals (e.g., the basket) and journals that publish IS research but also journals that we do not consider to constitute IS journals because they also publish non-IS research (e.g., *Communications of the ACM*, *Decision Sciences*). Nonetheless, this set makes for a reasonable comparison set because it includes a substantial portion of the population of IS research. We included all 51 journals in our analysis, but we report our results only for the 17 journals in Table 1 since we focus on them in this paper.

Table 8 presents the equation for PageRank. An edge of a directed network represents the citation relationship between journals, and the nodes indicate units that involve the citation behavior. In the present study, we set nodes as journals and the journal citation network has two kinds of edges: incoming edges and outgoing edges. A node's incoming edges (the number M in the equation) denote the number of citations the node has received, while the outgoing edges of a node denote the node's citing behavior (C in the equation).

Table 8. PageRank Equation

Equation	Symbols	
$\mathrm{PR}(p_i) = \frac{(1-d)}{N} + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{C(p_j)}$	N	The number of total nodes
	$PR(p_i)$	PageRank of node p_i
	d	Damping factor (a parameter), usually set to .85
	$M(p_i)$	Set of nodes that is linked to p_i
	$C(p_j)$	The number of links from p_j to other nodes

Table 9 presents the results for a journal's impact as a whole (without considering the number of papers the journal published). Thus, journals that published more papers (e.g., *DSS*) had more opportunity to have an impact than journals that published fewer papers (e.g., *MISQ*). This analysis shows that *MISQ* had the greatest impact of any journal in this set. In total, five basket journals placed in the top eight of the set, but, otherwise, one cannot easily conclude that they materially differed from the other journals in this set. No statistics to test for differences among page ranks exist because page ranks are neither independent nor normally distributed. For those interested in using a number to assess differences in PageRank, we offer Tukey's critical distance measure (calculated at $\alpha = .05$), which was 0.0114, although this measure assumes independence and a normal distribution, which our data did not have. The critical distance measure suggests that *MISQ* differed from *DSS* and that DSS differed from *EJIS*, but the remaining journals did not differ from those around them, although the bottom five differed from the top half.

Table 9 also shows the results for the impact that an individual paper has in each journal (i.e., considering the number of papers the journal publishes). This analysis shows that a paper in *MISQ* had the greatest impact. *MISQ* belonged to a class of its own, but we observed a much smaller difference among papers in other journals. Six of the basket journals placed in the top eight, but one cannot easily conclude they materially differed from the other journals in the set. Tukey's critical distance measure ($\alpha = .05$) was 0.0177, which suggests that a paper in *MISQ* differs from a paper in *ISR* but that papers in the other journals did not differ from papers in the journals around them. Papers in the top five journals differed from papers in journals in the bottom half.

[13] We are grateful to Paul Lowry for pointing out that literature reviews and theory building papers may buck this trend as they may cite a large number of papers in a detailed and critical manner.

Table 9. Journal-level PageRank and Paper-level PageRank

Journal level			Page level		
Rank	**Journal**	**PageRank**	**Rank**	**Journal**	**PageRank**
1	*MISQ*	0.0690	1	*MISQ*	0.1109
2	*DSS*	0.0573	2	*ISR*	0.0449
3	*EJIS*	0.0371	3	*DSS*	0.0380
4	*I&M*	0.0341	4	*JMIS*	0.0359
5	*JMIS*	0.0324	5	*EJIS*	0.0259
6	*ISR*	0.0288	6	*ISJ*	0.0238
7	*IJHCS*	0.0269	7	*IJHCS*	0.0186
8	*ISJ*	0.0268	8	*JSIS*	0.0186
9	*KBS*	0.0245	9	*JIT*	0.0147
10	*JSIS*	0.0234	10	*I&M*	0.0140
11	*IJIM*	0.0196	11	*IJEC*	0.0113
12	*JIT*	0.0173	12	*JAIS*	0.0101
13	*JAIS*	0.0137	13	*IJIM*	0.0066
14	*IJEC*	0.0122	14	*I&O*	0.0061
15	*ECRA*	0.0091	15	*ECRA*	0.0045
16	*EIS*	0.0062	16	*KBS*	0.0035
17	*I&O*	0.0049	17	*EIS*	0.0035

4.3 Summary

Here we briefly summarize our results from analyzing the validity and reliability of journal-level impact metrics.

- Traditional citation metrics rely on the arithmetic mean, and our analyses show that citation data are extremely skewed. In short, while a journal impact factor of 7.000 should imply a typical paper in the journal has attracted seven citations, the median figure for citations over two years was about four (or eight over five years).
- While median-based citation measures have greater validity than mean-based metrics, Figure 1 shows that only *MISQ* and *ISR* materially differed from the journals in our sample, which included six randomly selected journals from the Web of Science list.
- PageRank represents another metrics that does not rely on the arithmetic mean. This analysis shows the basket journals scored higher, although most of the differences were not significant. *MISQ* and perhaps *DSS* differed from the other journals in our sample. However, in the Web context in which researchers initially applied it, PageRank has received criticism for bias against new webpages that have not had time to receive citations. As a result, Google incorporated additional measures to incorporate page quality and future link potential to address this bias. In a similar fashion, PageRank would also be biased against newer papers that have had less time to accrue citations.

We make two main conclusions. First, the metrics above all focus on the journal level. Given that individual papers receive citations rather than journals per se, focusing so much attention at the higher level of granularity (i.e., the journal) may be inappropriate, and, as we show above, we can draw no definitive conclusions about any subset of IS journals beyond *MISQ* and perhaps *ISR*. Second, the metrics above all rely on citations as the metric of impact. However, advances in social media and technology have allowed a range of alternative metrics (altmetrics) beyond citations to emerge.

5 Altmetrics at the Paper Level

Altmetrics refer to an alternative set of emerging measures beyond the traditional ones such as citations. One can assess altmetrics at either the journal or paper level. Lin and Fenner (2013) point out that, while citations constitute an important way to measure impact, they represent a tiny fraction of user engagement with a paper. In studying more than 80,000 published papers, the Public Library of Science (PLOS)—an open access publisher that has been at the forefront in pioneering altmetrics—reported that 200 million paper page views were associated with 50 million PDF paper downloads (25%) but less than 500,000 citations (0.2%). Thus, it seems clear that the measures in previous sections, which all focus on citations, do not represent the full picture.

Paper-level measures (PLM) involve exactly what the name implies: measures that assess each paper individually. One can see from Table 9 above that paper-level analyses produce different rankings to journal-level analyses. PLOS suggests a PLM classification that reveals additional dimensions in which one can use a paper along a continuum from viewing/downloading a paper through to saving it to a reference manager, discussing and recommending it with tweets and likes, and eventually directly using it in research (i.e., a citation).

Each different measure focuses on a different part of the research process. Economists call some PLM leading indicators; that is, they typically occur before the phenomenon of interest occurs (O'Sullivan & Sheffrin, 2003). Leading indicators are correlated with the phenomenon of interest and may or may not be causal (O'Sullivan & Sheffrin, 2003). The phenomenon of interest is impact, which refers to a paper influencing the research thinking in another paper. In most cases, authors read a paper before it impacts their research. Typical reading measures include viewing it online or downloading it to read. Not all papers that are viewed/downloaded have an impact, and not all papers that have an impact are viewed/downloaded., but a possible correlation between the two measures exists in that views/downloads may predict citations in a later period (Brody, Harnad, & Carr, 2006; Guerrero-Bote & Moya-Anegon, 2014; Schloegl & Gorraiz, 2010).

Another indicator of impact (i.e., when a paper influences other research) occurs closer in time to when the impact occurs: discussing and recommending, which researchers do soon after a paper has influenced their research. There is a time lag, but the time lag to post to the Web is minimal. Once again, the relationship between these indicators and impact on research is correlational, not causal.

Researchers have long used citations to measure impact. Citations are lagging indicators—they appear after the activity of interest has occurred (i.e., the actual impact of someone's research) (O'Sullivan & Sheffrin, 2003). A paper attracts citations only after other authors have written and published their own work, an event that typically occurs years after the citing authors have engaged with the paper and it has influenced their research. The relationship here is also correlational because authors do not cite not all papers that influence their research and not all citations represent true impact (for instance, authors cite some research gratuitously or have research forced on them during the review process).

We build on the initial PLOS classification and summarize these categories in Table 10 in which we show indicative technologies that can help one generate these PLMs where available.

Table 10. Paper-level Measure Classification

Type	Category	Description	Relevant technology
Leading indicator	Viewed or downloaded	Users' accessing the paper online and, by implication, reading it	PLOS, AIS eLibrary
	Saved	Saving a paper to a citation manager	CiteULike, Mendeley
Direct indicator	Discussed	Mentioning a paper in shared comments on Twitter, in in-depth blog posts, in news articles, or in Wikipedia mentions	Twitter, Facebook, Google+, LinkedIn, Wikipedia, NatureBlogs, ResearchBlogging, PLOS Comments
	Recommended	Formally endorsing a paper via an online recommendation platform	F1000 Prime, ResearchGate
Lagging indicator	Citations	Formally citing a paper in a scientific forum	Google Scholar, CrossRef, PMC, Web of Science, Scopus
	Usages	Counting the number of citations in papers	
	Weighted Usages	Counting the number of citations in papers weighted by section in which they appear	

Some authors have argued that paper-quality assessment should move beyond citations to consider the number of times a paper receives a citation from another paper and in what section the latter cites the former (An, Kim, Kan, Chandrasekaran, & Song, 2017; Ding, Liu, Guo, & Cronin, 2013). In other words, they argue that, rather than considering a citation as a binary variable (0 or 1), one needs to consider it as a continuous variable. That is, we need to count the number of times a paper is used in each citing paper because one use (i.e., cited once) in a paper may indicate some usefulness, while a dozen or more uses in the same paper represents much more (e.g., that the paper had a core role in the citing authors' contribution). Thus, the usage count refers not to the number of papers that cite a paper but rather to the total number of times those papers cited the paper. Only two papers could cite a paper, but, if each one cited it 10 times, then the usage count would be 20. Such an approach requires one to analyze papers' full text, but, since one can access most papers in in full-text repositories, one need only use Google to search the Web to find them.

Likewise, some authors suggest that where a paper uses a citation has different implications for its value (An et al., 2017; Ding et al., 2013). After all, users actually interact with a paper's components. A citation in a paper's introduction may be a passing reference that adds little value. In contrast, a citation in the theory section suggests the citing paper's authors found the cited paper useful in developing new theory. Further, a citation in the method section suggests the citing paper's authors found the cited paper useful in conducting empirical research. In other words, a citation's placement may indicate that it had lesser or greater value to the various components in researchers' work depending on where they use it. Based on this line of thinking, one could weight usage counts by the section in which citations occur in addition to counting how many times the citing paper uses them. Once again, such an approach requires one to analyze papers' full text, but it is certainly technically feasible.

6 Conclusions

In this paper, we analyze 17 IS journals using several different measures for assessing research impact. Different measures produced different rankings, but, in general, we draw four key conclusions when we look across these analyses.

First, one journal stood out from the other IS journals. Across a variety of measures, *MISQ* usually had a higher impact than other IS journals, and the difference in magnitude was often material. Thus, we conclude that, among all IS journals, *MISQ* has the greatest impact.

Second, we cannot easily draw general conclusions for the other IS journals we examined. Different measures place the journals in different orders in terms of relative impact. No one other journal consistently placed higher or lower than the other journals. One or two measures suggest basket journals may have higher impact than other IS journals, but, for most measures, we found no differences that were both significant and meaningful between journals in the basket and journals not in the basket.

Third, the citation data was highly skewed as is common in most disciplines (Garfield, 2006). We found as much variation among papers in the same journal as among papers from different journals; thus, we found few systematic and stable differences between journals. Therefore, we conclude that one should not use the journal as an indicator of any specific paper's impact (and, by extension, quality).

Fourth, many researchers accept that one should not use the mean as a measure of central tendency with skewed data (Pagano & Gauvreau, 2000). The data above show that using the mean as a measure of central tendency significantly right-overstates the median (i.e., most likely outcome) as a measure of central tendency because a handful of outliers highly skew the mean. Thus, one should not use any measure (e.g., JIF) that uses the mean for citation data to measure research impact.

Our research has several important limitations. We studied only a small set of journals: the basket of eight (the putative eight "best" journals) plus nine others to augment our focus beyond the basket. Further, we studied only the last ten years of data since citation patterns can change over long periods of time. We studied only a small set of possible measures (the commonly used traditional measures plus two newer ones); however, other possible measures exist. And, of course, we studied only primarily IS journals, not general business or computer science journals that publish some IS research (e.g., *Management Science*, *Communications of the ACM*). Notwithstanding these limitations, we believe our study has several implications, which we discuss next.

6.1 Implications

We need to pause and consider our conclusions' implications due to their far-reaching and rather unpleasant nature. As scientists, we strive for validity and reliability in our measures (Campbell & Stanley, 1963; Cronbach & Meehl, 1955). We should expect the same of measures we use to assess the outputs of our labors.

First, our conclusions imply that the set of journal measures that we commonly use (e.g., JIF) lack accuracy because they use the mean as central tendency, not the median, and the mean and median point to two very different places as the central tendency. The JIF systematically distorts the underlying phenomenon, so we conclude that it fails the validity test. Thus, our conclusions imply that we should stop using the JIF (and other mean-based measures).

Second, as we look at these traditional biased measures or alternatives to them, such as PageRank, we found no clear pattern to indicate which journals are higher in quality than others; aside from *MISQ*, the other journals in the Senior Scholar's basket did not consistently significantly or meaningfully differ from the three close contenders for the basket (*DSS*, *I&M*, *I&O*) or even from randomly selected IS journals in the Web of Science list. This finding inescapably implies that the basket does not serve as a reliable measure of journal quality, which we find very awkward to write since we (the two primary authors) are AIS Senior Scholars with a long history of publishing in basket journals. One might wrongly interpret this implication as an attack on the institutions that we have helped build, but we do not intend it as such: we have followed the data where they lead and they show that the journal basket we helped build does not reliably serve as a measure of research impact. Thus, the first implication leads to the second: that we should stop using the journal basket.

The third and perhaps most challenging implication comes from the conclusion that one cannot reliably use a paper's journal to assess its quality. We regularly consider the number of papers in top journals as indicating research performance when we evaluate candidates for promotion and tenure (Cuellar et al., 2016b; Dennis, Valacich, Fuller, & Schneider, 2006). Informal discussions at any IS conference will reveal a preoccupation (a "fetishization" even) with certain journals, usually *MISQ* and *ISR*. IS researchers commonly state in their résumés that their "research has appeared in *MISQ* and *ISR* among other outlets". Our results show that using the journal that publishes a paper to indicate impact is a flawed criterion; it fails the reliability test. Thus, our findings suggest that we should no longer consider the journal as a reliable proxy for paper quality in promotion and tenure assessments.

6.2 Two Roads Diverging

Taken together, our implications suggest two different, diverging roads for the future (with apologies to Frost, 1916). First, they suggest we should continue doing what we already do now but better. Assessing quality at the journal level has an inherent appeal due to its simplicity. We could, for example, develop better journal-level measures that rely on the median to measure central tendency. Unfortunately, we believe that any attempt to "patch-up" the holes in our current approach would have problems; the papers in each journal simply vary too much in impact.

Even if we could develop measures that prove empirically better than our current measures at assessing journal quality, doing so misses the fundamental theoretical issue: journals do not create research value—authors and their papers do (Cuellar et al., 2016b). Journals primarily perform a curation function (Cuellar et al., 2016b) and may also add value during the review process. From a theoretical perspective, journals represent the wrong unit with which to measure research quality.

The other road involves moving beyond the journal and focusing on each paper individually. Researchers first created measures for journal impact in 1961 (Garfield, 2006) when one could not easily obtain paper-level data. Today, data are ubiquitous. Google Scholar updates citation data in real time—an unthinkable development in the 1960s when researchers developed the JIF and other commonly used measures.

We argue that we should use paper-level metrics going forward. By moving to paper-level measures rather than journal-level measures, research becomes more ecumenical and pluralist as we consider each research paper on its own merits, and we avoid artificially boosting or downgrading a paper based on the journal that publishes it. The assumption that the journal represents an automatic proxy for paper quality is a coarse-grained one. Paper-level measures are early or leading indicators and more fine-grained and peer informed.

At this point, we have a moderate amount of research and experience with the first road but little with the second road. At risk of pushing a metaphor too far, we note that Frost's (1916) narrator looked at the two roads and "took the one less traveled by". We do not have as much courage as Frost's narrator. Rather than immediately striking out on the less travelled road, we advocate more research and experimentation on the less travelled road so we can make a better informed decision about which road to take.

6.3 Closing Thoughts

We studied only IS journals, but we suspect that our conclusions apply to journals in other disciplines as well. As a discipline focused on information and systems, we have a responsibility to take the lead in tackling these issues.

Change takes time. Although we believe that IS as a discipline will eventually move away from traditional journal-level measures of impact as a proxy for quality, we do not believe we should abandon our traditional measures until we better understand the new measures. We envision a period of change as we develop and test new measures and more deeply understand their benefits and limitations.

Change is not easy. As Machiavelli (1532) notes:

> *There is nothing more difficult to take in hand, more perilous to conduct, or more uncertain in its success, than to take the lead in the introduction of a new order of things. Because the innovator has for enemies all those who have done well under the old conditions, and lukewarm defenders in those who may do well under the new.*

Nonetheless, we argue that we need to change given the lack of validity and reliability in our current measures of research impact.

Acknowledgements

We acknowledge the very helpful feedback and suggestions that Paul Lowry and John Mingers provided on early drafts of this work.

References

An, J., Kim, N., Kan, M.-Y., Chandrasekaran, M. K., & Song, M. (2017), Exploring characteristics of highly cited authors according to citation location and content. *Journal of the Association for Information Science and Technology*, *68*(8), 1975-1988.

Antelman, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, *65*(5), 372-382.

Bollen, J., Van de Sompel, H., Smith, J., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing & Management*, *41*(6), 1419-1440.

Bollen, J., & Van de Sompel, H. (2008), Usage impact factor: the effects of sample characteristics on usage-based impact metrics, *Journal of the American Society for Information Science and Technology*, *59*(1), 136-149.

Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual Web search engine. *Computer Networks*, *56*(18), 3825-3833.

Brody, T., Harnad, S., & Carr, L. (2006). Earlier Web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science and Technology*, 57(8), 1060-1072

Brzezinski, M. (2015). Power laws in citation distributions: Evidence from Scopus. *Scientometrics*, *100*(1), 213-228.

Campanario, J. M. (2011). Empirical study of journal impact factors obtained using the classical two-year citation window versus a five-year citation window. *Scientometrics*, *87*(1), 189-204.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago, IL: Rand McNally & Company.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Cuellar, M. J., Takeda, H., Vidgen, R., & Truex, D. (2016a). Ideational influence, connectedness, and venue representation: Making an assessment of scholarly capital. *Journal of the Association for Information Systems*, *17*(1), 1-28

Cuellar, M. J., Truex, D. P., & Takeda, H. (2016b). Can we trust journal rankings to assess article quality? In *Proceedings of the 22nd Americas Conference on Information Systems.*

Dennis, A. R., Valacich, J. S., Fuller, M. A., & Schneider, C. (2006). Research Standards for promotion and tenure in information systems. *MIS Quarterly*, *30*(1), 1-12

Ding, Y., Liu, X., Guo, C., & Cronin, B. (2013). The distribution of references across texts: Some implications for citation analysis. *Journal of Informetrics*, *7*(3), 583-592.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, *13*(4), e123.

Ferratt, T. W., Gorman, M. F., Kanet, J. J., & Salisbury, W. D. (2007). IS journal quality assessment using the author affiliation index. *Communications of the Association for Information Systems*, 19, 710-724.

Field, A. (2009). *Discovering statistics using SPSS.* Thousand Oaks, CA: Sage.

Fisher, J., Shanks, G., & Lamp, J. (2007). A ranking list for information systems journals. *Australasian Journal of Information Systems*, *14*(2), 5-18.

Frost, R. (1916). *Mountain interval.* New York, NY: Henry Holt & Company.

Gallivan, M. J. (2011). Are citations to academic journal articles a measure of quality or something else? An exploratory analysis with emphasis on design science and IS technical research. In *Proceedings of the Americas Conference on Information Systems.*

Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*, *295*(1), 90-93.

Garfield, E. (1979). IS citation analysis a legitimate evaluation tool. *Scientometrics, 1*(4), 359-375.

George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference, 17.0 update.* Boston, MA: Pearson.

Gillenson, M. A., & Stutz, J. D. (1991). Academic issues in MIS: Journals and books. *MIS Quarterly*, *15*(4), 447-452.

Glantzel, W., & Moed, H. (2002). Journal impact measures in bibliometric research. *Scientometrics*, *53*(2), 171-193.

Guerrero-Bote, V., & Moya-Anegon, F. (2014). Relationship between downloads and citations at journal and paper levels, and the influence of language. *Scientometrics*, *101*(2), 1043-1065.

Gravetter, F., & Wallnau, L. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Belmont, CA: Wadsworth.

Hamilton, S., & Ives, B. (1980). Communication of MIS research: An analysis of journal stratification. In *Proceedings of the International Conference on Information Systems.*

Hansson S. (1995). Impact factor as a misleading tool in evaluation of medical journals. *Lancet*, *346*, 906.

Hutchins, B., Xin, Y., Anerson, J., & Santangelo, G. (2016). Relative citation ratio (RCR): A new metric that uses citation rates to measure influence at the article level. *PLOS Biology*, *14*(9), e1002541.

Katerattanaku, P., & Han, B. (2003). Are European IS journals under-rated? An answer based on citation analysis. *European Journal of Information Systems*, *12*, 60-71.

Katerattanakul, P., Razi, M. A., Han, B. T., & Kam, H. (2005). Consistency and concern on IS journal rankings. *Journal of Information Technology Theory and Application*, *7*(2), 1-20.

King, M. M., Correll, S. J., Jacquet, J., Bergstrom, C. T., & West, J. D. (2016). *Men set their own cites high*: Gender and self-citation across fields and over time. *Sociological Research for a Dynamic World*, *3*, 1-22.

Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature, 411*(6837), 521.

Lewis, B. R., Templeton, G. F., & Luo, X. (2007). A scientometric investigation into the validity of IS journal quality measures. *Journal of the Association of Information Systems*, *8*(12), 619-633.

Lin, J., & Fenner, M. (2013). Altmetrics in evolution: Defining and redefining the ontology of article-level metrics. *Information Standards Quarterly*, *25*(2), 20-26.

Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D., Humpherys, S., Barlow, J. B., & Wilson, D. W. (2013). Evaluating journal quality and the assoclation for Information Systems Senior Scholars' journal basket via bibliometric measures: Do expert journal assessments add value? *MIS Quarterly*, *37*(4), 993-1012.

Machiavelli, N. (1532). *The prince*. Retrieved from https://www.gutenberg.org/files/1232/1232-h/1232-h.htm

Mingers, J., & Willmott, H. (2013). Taylorizing Business school research: On the "one best way" performative effects of journal ranking lists. *Human Relations*, *66*(8), 1051-1073.

Moed, H. F. (2007). The effect of "open access" upon citation impact: An analysis of ArXiv's condensed matter section. *Journal of the American Society for Information Science and Technology*, *58*(13), 2047-2054.

Murray M. (2014). Analysis of a scholarly social networking site: The case of the dormant user. In *Proceedings of the Seventeenth Annual Conference of the Southern Association for Information Systems.*

Opthof, T. (1997). Sense and nonsense about the impact factor. *Journal of Cardiovascular Research*, *33*(1), 1-7.

O'Sullivan, A., & Sheffrin, S.M. (2003). *Economics: Principles in action.* Upper Saddle River, NJ: Pearson.

Pagano, M., & Gauvreau, K. (2000). *Principles of biostatistics* (2nd ed.). Pacific Grove, CA: Duxbury Press.

Perneger, T. (2010). Citation analysis of identical consensus statements revealed journal-related bias. *Journal of Clinical Epidemiology*, *63*(6), 660-664.

Peffers, K., & Ya, T. (2003). Identifying and evaluating the universe of outlets for information systems research: Ranking the journals. *Journal of Information Technology Theory and Application*, 5(1), 63-84.

Polites, G. L., & Watson, R. T. (2009), Using social network analysis to analyze relationships among IS journals. *Journal of the Association for Information Systems*, *10*(8), 595-636.

Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact.* Cambridge, MA: The MIT Press.

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012a). *Altmetrics in the wild: Using social media to explore scholarly impact.* arXiv:1203.4745.

Priem, J., Groth, P., & Taraborelli, D. (2012b). The altmetrics collection. *PLoS One, 7*(11), e48753.

Rainer, R. K., & Miller, M. D. (2005). Examining differences across journal rankings. *Communications of the ACM*, *48*(2), 91-94

Rousseau, R. (1988). Citation distribution of pure mathematics journals. In L. Egghe & R. Rousseau (Eds), *Informetrics 87/88* (pp. 249-260). Amsterdam: Elsevier.

Saha, S., Saint, S., & Christakis, D. (2003), Impact factor: A valid measure of journal quality? *Journal of the Medical Library Association*, *91*(1), 42-46.

Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics*, *82*(3), 567-580.

Seglen, P. (1992). The skewness of science. *Journal of the American Society for Information Science*, 43(9), 628-638.

Singh, G., Haddad, K., & Chow, C. (2007). Are articles in “top” management journals necessarily of higher quality? *Journal of Management Inquiry*, *16*(4), 319-331.

Smart, J., & Elton, C. (1982). Consumption factor scores of psychology journals: Scientometric properties and qualitative implications. *Scientometrics*, *4*(5), 349-360.

Stewart, A., & Cotton, J. (2018). Does “evaluating journal quality and the association for information systems Senior Scholars journal basket…” support the basket with bibliometric measures? *AIS Transactions on Replication Research, 4*, 1-41.

Swanson, E. P. (2004). Publishing in the majors: A comparison of accounting, finance, management and marketing. *Contemporary Accounting Research*, *21*(1), 223-255.

Taraborelli, D. (2008). Soft peer review: Social software and distributed scientific evaluation. In *Proceedings of the 8th International Conference on the Design of Cooperative Systems.*

Trieschmann, J. S., Dennis, A. R., Northcraft, G., & Niemi, A. W. (2000). Serving multiple constituencies in the business school: MBA program vs. research performance. *Academy of Management Journal*, *43*(6), 1130-1141.

Valacich, J. S., Fuller, M. A., Schneider, C., & Dennis, A. R. (2006). Publication opportunities in premier business outlets: How level is the playing field? *Information Systems Research*, *17*(2), 107-125.

Vanday, J. (2012) Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*, *92*(2), 211-238.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), xiii-xxiii.

Willcocks, L., Whitely, E. A., & Avgerou, C. (2008). The ranking of top IS journals: A perspective from the London School of Economics. *European Journal of Information Systems*, *17*, 163-168.

Xiao, B., Cheung, C., & Thadani, D. (2011). Assessing the quality and knowledge contribution of top IS journals: A comparative citation analysis. In *Proceedings of the European Conference on Information Systems.*

Zahedi, Z., Costas, R. & Wouters, P. (2014). Assessing the impact of the publications read by the different Mendeley users: Is there any different pattern among users? In *Proceedings of the 35th International Association of Technological University Libraries.*

About the Authors

Brian Fitzgerald is Director of Lero—the Irish Software Research Centre, where he previously held the role of Chief Scientist. Prior to that, he served as Vice-President Research at the University of Limerick. He also holds an endowed professorship, the Krehbiel Chair in Innovation in Business & Technology, at the University of Limerick. He holds a PhD from the University of London. His research interests lie primarily in software development, encompassing open source and inner source, crowdsourcing software development, agile and lean software development, and global software development. His publications include 15 books and over 150 peer-reviewed papers in leading international journals and conferences in both the information systems and software engineering fields, such as *MIS Quarterly*, *Information Systems Research*, *IEEE Transactions on Software Engineering*, and *ACM Transactions on Software Engineering Methodology*. Prior to taking up an academic position, he worked in the software industry for about 12 years in various sectors (including finance, telecommunications, manufacturing, bespoke software development) and countries (Ireland, Belgium, Germany).

Alan R. Dennis is Professor of Information Systems and holds the John T. Chambers Chair of Internet Systems in the Kelley School of Business at Indiana University. He was named a Fellow of the Association for Information Systems in 2012. Professor Dennis has written four books and more than 150 research papers in journals and conferences. His research focuses on three main themes: team collaboration; fake news on social media; and information security. His research has been reported in the popular press over 500 times, including the *Wall Street Journal*, *USA Today*, *The Atlantic*, CBS, PBS, Canada's CBC and CTV, UK's *Daily Mail* and the *Telegrap*h, Australia's ABC, France's *Le Figaro*, South Africa's *Sowetan Live*, Chile's *El Mercurio*, *China Daily*, India's *Hindustan Times*, and Indonesia's *Tribune News*. He is the co-Editor-in-Chief of *AIS Transactions on Replication Research* and the President of the Association for Information Systems.

Juyoung An is a master's student in the Text and Social Media Mining Lab of the Department of Library and Information Science at Yonsei University. She has Master in Information Science at Indiana University. Her areas of interests are Text Mining, Natural Language Processing and Deep Learning. She is working as an AI researcher to develop various NLP engines in Korea.

Satoshi Tsutsui is a doctoral candidate at the IU Computer Vision Lab in the School of School of Informatics, Computing and Engineering at Indiana University. He is interested in data science (broadly defined) and computer vision. He has worked on scientific literature mining, image captioning, object detection, and semantic segmentation.

Rishikesh C. Muchhala is a Chicago-based management consultant with PricewaterhouseCoopers LLP in the Health Industries business. He holds a Master of Science in information systems from the Kelley School of Business at Indiana University Bloomington. While he dons a practitioner's hat in his day job, his research interests include societal impact of technology, adoption of information systems, misuse, underuse, and wastage of technology.